DOCUMENT RESUME

ED 046 863 SP 004 547

AUTHOR	Ornstein, Allen C.
TITLE	Methods for Conducting Teacher Behavior Research: With Implications for Teachers of the Disadvantaged.
PUB DATE	70
NOTE	35p.
EDRS PRICE	EDRS Price MF-$0.65 HC-$3.29
DESCRIPTORS	Academic Achievement, Classroom Observation Techniques, Disadvantaged Youth, *Educational Research, Observation, *Performance Criteria, Personality Tests, Student Behavior, *Teacher Behavior, *Teacher Qualifications, *Urban Education

ABSTRACT

In this paper on research into teacher behavior, the author considers first the current status of such research and the fact that the problem is so complex that no one knows or agrees what a competent teacher is. The difficulties of identifying and measuring teacher behaviors are discussed. The four broad areas of measurement (observation, student behavior and achievement, ratings based on recall, and personality tests) are considered in detail, together with additional problems including the noncumulative nature of earlier research; the problem of distinguishing, control and analysis of the variables; the rapid sequence of classroom events; the non-scientific nature of teaching; and teachers' lack of enthusiasm for researchers, particularly in inner-city situations. In an attempt to improve this situation, twenty-seven recommendations for future research are listed. The paper is based on Chapter 3 of the author's unpublished dcctoral dissertation, Selected Teacher Behavior Attributes Rated as Desirable by Ninth-Grade Disadvantaged Students and Ninth-Grade Teachers of the Disadvantaged. (ED 043 563 and ED 043 564 are related documents. (MBM)

ED046863

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
OFFICE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY.

METHODS FOR CONDUCTING TEACHER BEHAVIOR RESEARCH: WITH IMPLICATIONS FOR EDUCATING THE DISADVANTAGED[1,2]

Allan C. Ornstein

The Status of Research on Teacher Behavior

The research on teacher behavior is voluminous and contradictory. The problem is so complex that no one knows or agrees upon what a competent teacher is. Broudy (1969) contends that we "can define good teaching any way we like [p.583]." If we are unable to define a good teacher, it follows that it is injudicious to formulate scientifically based generalizations about good teacher behavior.

Also, we lack the ability to evaluate the symbolic aspects of teaching, and we lack sufficient knowledge about learning to evaluate teacher behavior or instruction adequately. Goheen (1966) points out that teacher behavior cannot be defined and analyzed, and therefore, "there will always be teachers who will break all the rules and yet be profoundly successful [p.221]." No single teacher can possibly possess all the traits listed in several studies; moreover, the research findings generally are not applicable to a specific classroom situation.

---

[1] "This paper is primarily based on the author's first draft of the "Conceptual Framework," in Selected Teacher Behavior Attributes Rated as Desirable by Ninth-Grade Disadvantaged Students and Ninth-Grade Teachers of the Disadvantaged. Unpublished doctoral dissertation. New York University, 1970.

For purposes of documentation, the aforementioned chapter should be read.

[2] Since research on teacher behavior for the disadvantaged is almost nonexistent, we are required to examine the broad field of research on teacher behavior. It is assumed that although the concepts - definitions, measurements and procedures vary with a specific study, they are general in nature and implementation, and therefore germane to all teachers - including those who work with the disadvantaged.

SP004547

Not only has teacher-behavior research failed to produce worthwhile results, but also the findings are either meaningless or else simply confirm "common sense." The results up to now are small in proportion to the outlay in time and effort; the field is becoming unmanageable; the investigators themselves do not know what to make of their findings. Berelson and Steiner (1964) summarized and dismissed research on teacher behavior in five words - "there are no clear conclusions [p. 441]." Why? The problem mainly involves the "criteria," whereby the investigator identifies and measures teacher behavior. This paper will be primarily concerned with this twofold problem .

Identifying Teacher Behaviors

Among the voluminous amount of research on teacher behavior, there are many options for choosing teacher-behavior characteristics; this in itself causes a problem. Our inability to define or agree upon the behaviors which constitute "good" teacher behavior or "effective" teaching has confused researchers and/or caused inconsistencies among the research findings. A particular pattern of teacher behavior cannot be advocated or reduced to a formula or rule. The reasons are sevenfold.

First, some investigators contend - either directly or indirectly - that it is fruitless to try to identify "good" teacher behavior, because teaching involves an interpersonal relationship between teacher and student(s) which must be described and analyzed. Nevertheless, these investigators fail to provide an empirical method for evaluating their recommendations or for conducting research.

Second, confusion over a variety of terms, such as "teacher personality," "teacher competence," "teacher performance," "teacher behavior," "teacher-behavior characteristics," "teacher traits," etc. adds to the general problem. Even worse, the definition and usage of these terms vary among different researchers.

Third, there are too many teacher behaviors to analyze or assess, and there is lack of agreement upon a common method for evaluating teacher behavior. In addition, there are no clear or acceptable methods for categorizing and/or identifying teacher behaviors. The teacher-behavior categories are vague and ill-defined. There is difficulty in classifying teacher behavors into proper and valid dimensions; teacher behavior from one study often cannot be categorized into the same dimension in another study. On the other hand, different teacher behaviors categorized into a specific dimension, despite their "independence," are often related either logically or statistically. The validity or "independence" of teacher behaviors which are categorized into dimensions is likely to decrease with the increase of dimensions while mutual exclusiveness decreases. Yet, if the teacher-behavior dimensions are decreased, the findings often are oversimplified and little worthwhile data are forthcoming. Thus, it is questionable whether a set of criteria can be developed to provide sufficient properties for classifying teacher behaviors.

Fourth, there are too many "similarities" and "dissimilarities" among the different teacher-behavior categories, causing serious and confusing analytical problems, making it nearly impossible to determine the differences within a teacher-behavior classification. For example, a teacher who "gives direction" would be considered as exhibiting "direct behavior" by Flanders (1965), "controlling behavior" according to Hughes (1962, 1965) and Ornstein (1970), "routine behavior" according to Gallaher and Aschner (1963), "directing and managing behavior" according to Meux and Smith (1964), and "responsible behavior" according to Ryans (1960). These different teacher-behavior categories, although somewhat similar, tend to invalidate comparisons between different studies. A particular teacher behavior judged to be "effective"

in one study can be judged "ineffective" in another study. The only consistencies are the obvious teacher behaviors; for example, "friendly" behavior is indicative of a "good" teacher and "hostile" (opposite-type) behavior is indicative of a "poor" teacher.

Fifth, there are thousands of descriptive words that may be applied for describing and classifying teacher behavior. For example, dealing with one teacher behavior alone, namely, verbal behavior, Flanders (1965) employed 7 different examples while Zohorok (1968) used 175 different examples. Assuming content validity was established in both cases, who is right, and who determines who is right? Judges are biased, so is the reader. Into how many different components can verbal behavior, for that matter any type of teacher behavior, be subdivided? No one really seems to know, or at least agree. Linguistic usage, confusion over words, and/or interchangeability of words cause difficulties concerning agreement on operational or behavioral meanings of teacher-behavior categories, or, in the way in which teacher behavior occurs, as well as the nature and scope of the behavior. For example, this author (1970) used "welcomes and is respectiful of views other than own" as a behavior phrase to help describe Affective Teacher Behavior. A similar teacher behavior, "sincere sympathy with a pupil's viewpoint [p. 88]," is categorized by Ryans (1960) under Understanding Behavior. Dumas (1966) ranked "sympathy with pupil viewpoint [p. 24]" with Empathy. Medley and Mitzel (1963) identified "tried to see pupil point of view [p. 276]" with Teacher Climate. Remmers (1963), reviewing different rating scales, reported "accepted students' viewpoint with open mind [p.342]" under Adequacy of Relations with Students. Sontag (1968) itemized "shows interest in the viewpoint of pupils [p. 395]" with Concern for Students. Jersild (1940)

linked "permitted expression of opinion [p. 144]" with Teacher Performance. This type of discrepancy, this inability to agree upon operational terms, causes the findings to be inconclusive and ambigious and the research and related literature to be misleading and ungeneralizable.

Sixth, even when there is agreement on "good" teacher behavior, it is wrong to assume that there is a commonly agreed upon meaning regarding the words used to describe such behavior. Teacher-behavior concepts and definitions have different meanings within different groups or subjects i.e., students, teachers, supervisors (in part, because of their different roles) and even within the same group of subjects. This problem is also evident among the various investigators themselves, even though they often attempt some kind of acceptable validity. For example, this author (1970) used 14 items for classifying cognitive teacher behavior. Gallagher and Aschner (1963) organized the same teacher behavior into 4 dimensions, based on the Guilford (1954) model of intellect, along with 11 subdimensions and 14 items. Masia (1965) organized cognitive teacher behavior into 6 dimensions, based on the Bloom et al (1956) taxonomy of educational objectives, along with 17 items to illustrate the 6 dimensions. Warren (1968) presented 40 items to evaluate the same teacher behavior. Not only do almost all the specific items differ among these investigators, but the catagorization and validity process becomes incongruous when it is pointed out that, with the exception of Gallagher and Aschner, the other investigators solely (Masia and Warren) or largely (Ornstein) refer to Bloom for purposes of defining cognitive teacher behavior; their definitions being similar. It seems that the investigators have their own vocabulary for defining specific teacher behaviors.

Seventh, judgments about teacher behavior are socially biased. Teacher behavior varies with the nature of goals; most studies fail to take this into

account and are therefore misleading. Also, teacher behaviors involve values and social outcomes which cannot be quantified.

Measuring Teacher Behaviors

In addition to being unable to agree on an acceptable or valid list of teacher behaviors, there is a lack of agreement on how to measure teacher behavior. Furthermore, the methods for measuring teacher behavior often seem questionable in terms of reliability and/or validity. At best, the reliability and validity of measurements of teacher behavior should be considered only relative to a defined situation, which in turn, yields relatively ungeneralizable findings. With this, let us proceed to discuss methods for assessing and/or correlating teacher behavior; they fall into four broad areas: (1) observations, (2) student behavior and achievement, (3) tests based on recall, and (4) psychological tests.

Observations. Observations may be classified into three types: (1) post-session evaluation, whereby after the class session is finished the observer makes broad evaluations of what went on, (2) sign observation, whereby the observer rates a specific list of behaviors by some specific unit of time, i. e., "moves," "acts," etc., and (3) categorical observation, whereby the observer uses a scale to rate a specific list of teacher behaviors. In general, all three techniques are somewhat biased and deficient. The first method, however, is the least reliable and valued, but is the most popular one used by supervisors to rate teachers. The supervisor usually observes a fraction of what goes on in the classroom and remembers only a fraction of what he observes. There are six basic problems related to observing teacher behavior.

First, observations of teacher behavior are limited because of the small number of observations upon which a rating is based. Teacher behavior changes daily, and observers should rate teachers over a period of time and on many

separate occasions. No matter how reliable or valid the observer's assessment procedure, the results tend to be distorted, for the teacher tends to put on an act while being observed. Operating, similarly, is what is called "demand characteristics;" the subject (teacher) is willing to cooperate as indicated by the fact that he has usually consented (however, sometimes by coercion) to the observer's presence, and therefore perceives an "acceptable" role, which, in turn, changes his behavior. The presence of the investigator (observer) creates what is called the "Hawthorne effect" - novelty, awareness of participation, and/or an altered situation - on the subject (s), and these effects are often too complex to determine. The investigator (again, observer) transmits what is termed "bias effect," that is, his own hunches or prejudices which are often one of the factors which prompted the study; they are transmitted to the subject(s) and to those with who they are interacting (teachers and students) in such a way as to alter the subject(s)' behavior.

Second, observers are influenced by their own values and role interpretation of what constitutes a "good" teacher. Even the age and sex of the observer and teacher influences the rating. The cues upon which the observer bases his judgments vary in importance from observer to observer, and even with the same observer, for different teachers. Each of us has a preferred set of teacher behaviors, and even though the observer has a specific list to interpret, he tends to concentrate on the favored items and to bypass the others. The observer is subject to the "halo effect," whereby he rates the teacher's behavior in the direction of his general impression of the teacher. Guilford (1954)[1] affirms that the observer's rating are also distorted by the

[1]Actually, Guilford is specifically referring to raters, not observers; however, in effect, the observers are rating teacher behavior with some kind of rating scale. Thus Guilford's discussion is germane to observers.

following: (1) "error of leniency," a tendency of the rater to rate low or high, no matter what the reason; (2) "error of central tendency," whereby the rater (observer) is reluctant to make extreme judgments about others (teachers); (3) "constant error," whereby the rater tends to rate others in the opposite direction of his own behavior - for example, the observer who is businesslike tends to perceive the teacher as less businesslike, or the observer who is not too businesslike tends to perceive the teacher as more businesslike.

Third, the observers not only are biased, but they often lack real knowledge about the specific classroom problems which are affecting the teacher's behavior. The same teacher behavior means different things to the students than to the observers. Observing teacher behavior is sometimes of little value, for the basic referents of effective teaching are linked with the teacher's personality.

Fourth, there is the problem of "observer loading," namely that it is "humanly impossible" to objectively observe, or just observe, all teacher behavior or classroom phonomena, and it is unlikely, too, that the phonomena which are observed will be given appropriate weights. The observer cannot rate all teachers equally well on all traits. Usually overt behavior is measured by the observer and nonverbal and/or intangible entities pertinent to teaching are often overlooked by the observer, i.e., an expression or glance that is easily understood by the students. Even though the teacher says the "right words" or behaves in the "right way," his "real" attitude is evident to the students (but often overlooked by the observer), which in turn, affects the classroom process. The direction of verbal communication - who talks to whom, and whether the teacher's statement is directed at an individual student as such or at an individual student as a member of a clique or class, is important but difficult for the observer to discern.

Fifth, most observers cannot state precisely the reasons for their judgments. Observers are unable to observe teachers systematically. They often lack sufficient time to make their evaluations. Their intelligence, according to Guilford (1954), accounts for as much as 15 percent of the score variance of a measured teacher behavior. No matter how well trained, most observers have difficulty in distinguishing one teacher behavior from another as belonging to a specific class of behaviors, and once observers have been trained, it cannot be assured that their reliability will remain high over a period of time.

Sixth, although it might be pointed out that some of the above problems can be reduced with the introduction of visual and/or auditory tapes, the noise level of the class, the mechanical problems, and the cost of recording do not make the mechanized approach as valuable as it might seem. Also, the recordings are produced by the observer or filtered through his eyes, and therefore, they still incorporate and reflect most of the above contaminating factors.

Student behavior and achievement. Practical methods for evaluating teacher as a function of student behavior have not yet been developed; five reasons are set forth. First, a major problem seems to be that different student behaviors are assessed with different teacher behaviors, making it difficult to obtain a consistent thread or relationship. Three examples should indicate the infinite number of combinations. Ryans (1960) assessed 4 student behaviors (alertness, responsibility, confidence, and initiative) with 18 teacher behaviors or 3 broad teacher patterns (warm, understanding, friendly; responsible, business-like, systematic; and stimulating, imaginative, surgent). Perkins (1964) assessed 9 student behaviors which, for the sake of brevity, may be subsummed under two categories (work activity and social activity) and 10 teacher behaviors - along a supportive-nonsupportive continuum. Harvey, et al. (1968)

assessed 7 student behaviors (cooperativeness, involvement, activity, nurturance seeking, achievement, helplessness, and concreteness of responses) with 3 teacher behaviors (resourcefulness, dictatorialness, and punitiveness.) It should be pointed that not one student or teacher behavior is the same, and even assuming there were two alike, the definitions would probably differ. In short, dissimilarity of student and teacher behaviors, along with differences in definitions, make comparisons extremely difficult and often misleading.

Second, student behavior seems functional to countless other variables, which are often uncontrollable and too multidimensional to analyze effectively. A few examples suffice: subject matter, peer-group relations, classroom activities, school conditions or school norms, and community relations. Examining the subtleties of just one variable - subject matter - with regard to student behavior should make the reader more aware of the immense problem involved in coping with all the known variables for purposes of evaluating teacher behavior. Some subjects call for energetic, active student behavior, e.g., music, drama, and physical education; others usually demand a monotonous, quite drill-like atmosphere, e.g., mathematics and foreign language. The question is whether the investigator realizes, considers, and adjusts his analysis to the different reasons for the different atmospheres. To the knowledge of this author, the answer seems to be no; therefore, the research on teacher behavior tends to show that mathematics and foreign language teachers are more authoritarian, businesslike, and/or responsible than music and English teachers, and no qualifying explanation is provided.

Third, assuming the investigator is aware of the many variables, how does he evaluate them in regard to student behavior. For example, the teacher asks a question, but no answers are forthcoming. Is this because the observer is in the

room? Is it because the students are bored or confused? How many students are hungry, or haven't had breakfast? Are the students reacting to their present teacher or their previous one? Is the weather or lighting[1] influencing the students? Does the day of the week or time of the day influence the students? Is student behavior the same on Monday morning and Friday afternoon? Is there a basketball game scheduled for the afternoon? Rarely does the research consider these subtle factors in their assessment of student or teacher behavior and to dismiss them as being insignificant or minor is a mistake, for they comprise a large part of the classroom situation. It can be argued that randomization should solve this problem; however, it cannot be guaranteed that the biases will continue to operate in the same direction or that the sample will be large enough.

Fourth, observers have difficulty in distinguishing between teacher-behavior intent and effect on student behavior. There is no agreement on what constitutes desirable student behavior. By the same token, it is possible for the teacher

---

[1]Recently, an experienced teacher argued that if two elementary school classes, A and B, were matched according to I.Q. and reading achievement, then pre- and post-tested after one year, and if the students in class A were reading one year higher than the students in class B, it would be safe to assume that teacher A was a "better" teacher. The author remarked that there were still too many variables, many unidentifiable or too nebulous, to warrant such a conclusion. The students in class A could have been in a room for the whole year in which the sun shined through the windows, whereas the students in class B might have been situated on the dark side of the school. Would the investigator, teacher, or examiner note the difference? Do we look for such small, subtle differences? How important are these differences? No one seems to mention such variables, which does not necessarily mean they are irrelevant. Perhaps the amount of light or the difference in the amount of light in a classroom is a key factor, which we tend to ignore. We don't really know; there are no studies, it seems, that say otherwise. The teacher contended the author was being absurd. "No, I'm reflecting the absurdity of research on teacher behavior." Total all the so-called minor and unidentifiable variables, include, too, the variables we recognize, but cannot agree upon with regard to definition and weighting, and we have no firm conclusions.

to know the rules of "good" behavior or how to implement "good" teacher behavior (assuming we could agree on what "good" teacher behavior is) without this necessarily guaranteeing "good" student behavior. Also, there seems to be an intangible social relationship between teacher and students, which affects students' behavior, but which cannot be prescribed or defined.

Fifth, one might question how the observer goes about observing the students' behavior. If the observer is in the rear of the room, which tends to be the usual method, he perceives the back and external part of the cranium, not the students' face and telling gestures. Surely, the observer must miss some behavior, perhaps even be deceived. If the observer positions himself on the side or front of the room, the "Hawthorne effect" and "observer's biases," as previously defined, are probably enhanced. Finally, since the assessment of student behavior usually involves the observation of students, the problems of observation as previously described, too, are generally applicable for assessing student behavior.

Limitations are apparent in using student achievement as a criterion for assessing teacher behavior; they are divided into five points.

First, learning principles are vague in relation to the actual classroom process. Student achievement is often incidental or is inconsistent with "good" teacher behavior. It is difficult to distinguish which modes of teacher behavior - the verbal interchange, general strategies, reinforcement techniques, etc. - are related to student behavior.

Second, it is difficult to equate the effects of a particular teacher with student achievement. The variations in student personality, intelligence, past achievement, and/or environmental factors make it difficult to objectively measure student achievement as a function of teacher behavior. Then, there are other contaminating factors, such as mass media, low or high pretest scores,

time interval between the reliability tests and retests, school conditions, etc. The initial and final achievement tests are usually administered in a relatively short time interval; therefore, the magnitude of differences between tests tend to be small. In this connection, it is difficult to obtain reliable scores when the magnitude of differences is small.

Third, achievement tests only measure a small portion of the desired change expected of students as a result of the teaching-learning process. It omits, for example, the personal and social growth of the student.

Fourth, many educators contend that achievement tests often lack acceptable reliability and validity scores. Achievement tests are considered to be culturally biased, and they tend to discriminate against creative and/or intelligent students; they often lead to erroneous connotations and negative "dysfunctional" outcomes such as the "self-fulfilling prophecy."

Fifth, if an investigator or an observer is in the classroom, and especially if he is administering a student achievement test, many of the problems concerning the "Hawthorne effect" and "bias effect," as previously mentioned, appear to operate with the students, too.

Ratings based on recall. Teacher-behavior ratings based on recall are generally made by supervisors, teachers, and/or students. Studies of what constitutes "good" teacher behavior show that while evaluations made by supervisors, teachers, and students may be consistent they are often contradictory, or not significantly related.

There tends to be substantial agreement that students are the most worthwhile and honest raters of teacher behavior; moreover, they appear to be reliable raters. Remmers (1963) affirms that as long as 25 or more students' ratings of teacher behavior are used there is considerable reliability. Beyond

their assumed reliability, the students are considered more valuable as raters because they see the teacher perform on many occasions and under varied conditions. In fact, students within the same class can be regarded as many observers rating one teacher; they are considered good judges because as a group they represent a constant variable. Since the students' feelings are a major factor in determining the classroom climate, they are qualified to rate teacher behavior.

Nevertheless, teacher behavior ratings which are based on student recall do have limitations. It is contended that students lack knowledge of what is "good" teacher behavior; they are immature, their ratings are influenced by how easy the teacher is, and their ratings negatively affect teacher morale. Teachers affect students in different ways, and what accounts for these differences is not so much the teacher's behavior but the students' personalities. In this connection, then, the teacher can employ "good" teacher behavior but be rated as a "poor" teacher, because the students' rating reflect their attitudes and values. Perception of teacher behavior sometimes varies with student-achievement level or with interaction between students and teachers. Student ratings, while reliable, can vary from grade to grade. Low student-reliability scores for specific teacher-behavior items may be ignored so long as the broad teacher-behavior dimension is generally acceptable.

Referring, now, to the human rater in general, no matter if he is a consultant supervisor, teacher or student, the problems listed by Guilford (1954) under the discussion of observation - "halo effect," "error of leniency," "constant error," "error of central tendency," - tend to affect raters who are assessing teacher behavior on recall, too. Other factors that tend to affect raters, according to Guilford are: (1) sex, (2) age, (3) intelligence, (4) understanding of directions,

(5) understanding of purposes, (6) sufficient time to complete the ratings, (7) possession of the traits being measured, (8) different criteria raters employ for assessing the same trait or teacher behavior.

Second, when dealing with items about personality or behavior, raters often give answers they perceive as right to the investigator or test examiner. If the test examiner is not perceived as a member of the raters' "reference group," they sometimes give "socially acceptable answers." There are other contaminating factors, such as the way the raters perceive the test examiner's clothing, socio-economic class, race, and name. Attitudes are difficult to measure, because raters sometimes have two different attitudes - one for friends and relatives, the other for formal surveys. Raters are not always motivated or honest, and findings often reflect their lack of information concerning the "desirability" or "undesirability" of what is being measured.

Third, raters, because they are human, are "imperfectly reliable" and their judgments are susceptible to selective perception and memory, as well as lack of sophistication as to what is important or how to complete the inventories. Test interpretation of teacher behavior vary according to raters; moreover, the scale values (assuming the investigator weighs each item) are determined by judges or raters judging the location of each response in terms of "desirability," "favorability," "importance," etc., and their attitudes are biased, thus influencing the subsequent scoring of each respondent. Finally, the problem of what is acceptable reliability and validity seems germane to the construction of rating scales.

Personality tests. The problems centered around personality tests can be divided into four areas. First, psychologists are unable to agree upon a definition of personality or specific personality traits; therefore, it seems

that data provided by one instrument does not necessarily yield analogous data, even though the findings may correspond. Psychological tests are restricted because we do not know which personality traits are predictive of "good" teacher behavior. Some investigators point out that not enough is known about personality traits to predict teacher behavior, while others affirm that psychological tests are limited because teacher behaviors are inadequately defined. Getzels and Jackson (1963) conclude that "very little is known for certain about the nature of teacher personality, or about the relation between personality and teacher effectiveness [p. 574]."

Second, most psychological tests have uncertain validity. It has been shown that psychological tests administered to teachers can be biased in favor of extreme responses. Signing the answer sheet of a psychological test makes an important difference in the teachers' answers. Psychological tests are susceptible to "faking" by teachers. As Medley (1961) indicates, "You can't believe the answers teachers give ... [moreover], those who know how to get along with pupils also know how to get along on personality tests as long as they are not too subtle [p. 153]."

Third, it is inappropriate to rate someone on some point along some continuum because personality traits are abstract. "Unsure" or "neutral" responses on personality or attitude tests tend to make the instrument less valid. "Absolute" responses often force the respondent to make an unwarranted decision. It is impossible to control all the teacher-behavior variables - sex, intelligence, age, education, experience, etc. - and correlate psychological scores and teacher behavior.

Fourth, personality tests are usually given to a group of teachers, without distinguishing differences in subject field, grade level, or education, thus neutralizing differences that may exist. Most psychological tests are

evaluated in terms of teacher success or teacher behavior, as if an "ideal" existed. Teacher personality scores are often oversimplified, and do not really help to describe the effect of the teacher's personality upon the student or class, the teacher-student interaction - what teaching involves.

In short, teacher-personality scores tend to be useless, contradictory, and lacking in psychological and "common sense." Also, many of the general problems of and limitations of test construction and evaluation, especially those concerning reliability and validity are relevant to personality tests.

Some Additional Problems Related to Research on Teacher Behavior

First, much of the previous research on teacher behavior is nonculmulative, in the sense that researchers measure different phenomena and variables, use different terms, methods and assumptions. When findings are not in harmony with existing data, it behooves the researcher to explore further into his results and examine the reason; however, ideas and findings on teacher-behavior research are usually promulgated without much reference to, and with apparent disregard of, what others say or report. Indeed, researchers lacking a common framework to work with.

Second, the problem of distinguishing, controlling, and analyzing what seems to be an endless number of variables may be too difficult, if not an impossible problem to cope with, for obtaining worthwhile data.[1]

---

[1]A few examples of teacher behavior variables are the time, place, school morale, school goals, teacher training, sex, age, grade level, type of classroom, community, etc. Combine this list with an endless list of student, teacher and enviornmental variables as well as unidentifiable variables as previously defined.

At best, the findings of a study on teacher behavior should be considered relative to the variables being manipulated - bearing in mind that many haven't been identified. For example, does a "friendly" teacher, which usually connotes a "good" teacher, have the same effect all the time, in all schools, no matter what his age or sex, no matter what grade level or subject, in the classroom as well as when conducting student traffic in the cafeteria, hallways or auditorium, and with all types of students? Are there shades of differences or major differences, and to what extent, with which variables?

There is simply no adequate criterion and/or list of variables against which a list of teacher behaviors can be validated and/or compared.

Third, classroom events occur at such a rapid pace, involving hundreds of interpersonal changes per hour; it cannot be accurately systematized into a scheme that can help teachers in their actual situation. On the other hand, insignificant variables are often manipulated into research, producing trivia. Many variables are unpredictable. Individual variables are sometimes "coupled" - meaning that each affect the other, combining into new components; moreover, the new components affect other variables and cause still other components, some of which cannot be thoroughly distinguished or measured. Similarly, two or more variables do not necessarily reflect causal relationships, or reflect what they seem to show, but may reflect other variables which act upon the ones that seem more evident. Variables are multidimensional, not linear, and therefore, confusing and difficult to assess. In short, there are an infinite number of variables - whose interaction and importance are relative, some of which are unidentifiable, uncommon, unique, and unpredictable, but important although we are not sure to what extent - yielding uncontrollable data, making assessment of teacher behavior extremely difficult.

Fourth, analysis of teacher-behavior research may be beyond scientific analysis, because the act of teaching, itself, may be unscientific. The process of teaching may not be a natural phenomenon that is suitable or controllable for scientific inquiry. Teacher behavior is difficult to assess, perhaps, because the act of teaching involves working with complex organisms. Teaching is novel, but absolute in the sense that all teacher behavior and teaching situations are new, making it obscure for researchers. As previously mentioned, teaching is too complex for an abstract or scientific description. As

previously mentioned, many teaching acts, especially nonverbal ones, go unnoticed, or are difficult to make sense out of and evaluate. It is possible that teaching cannot be quantified into global or recognizable terms. As of now, a technical language has not yet been developed, one that is empirically based, to evaluate the actual teaching phenemona.

Fifth, researchers are using, for their analysis of teacher behavior and teaching, behavioral-science approaches, not educational theory or approaches applicable for classroom analysis. Psychology is a science and teaching is an art, or, at least, teaching is both a science and an art, and psychological approaches or behavioral approaches cannot fully describe an art, or, at least, teaching cannot be completely analyzed by scientific methods. Teaching often depends on feelings, hunches, and/or insights, and those are often more useful "to determine what and how" to teach than are scientific findings.

Sixth, the fact that teachers generally seek practical, "cook-book" approaches, although often condemned by other educators as mechanical or hazardous, suggests that teacher-behavior research is meaningless for teachers; it may suggest that scientific analysis and/or theoretical formulations are inapplicable to the classroom, too. While teacher-behavior research can be formulated, many feel it is unrelated to practice, to the actual classroom situation, because every teaching situation is different. The teacher who attempts to apply research findings to his classroom may not obtain similar or expected results, even though the teacher carried out the behaviors which are suggested by the research. The best advise (which may be based on research) can sometimes be harmful, because each teacher, student, and group of students - each situation - varies.

Seventh, the researcher often fails to put himself in the position of the teacher, and therefore, omits relevant facets of teacher behavior. Teachers use vague terminology to define their own classroom behavior and are unable to systematize or explain what they are doing. Likewise, many teachers know what they are doing in the classroom, but are unable to specifically state it in precise terms. Teachers and researchers use different terms to describe the same teacher behavior; moreover, the researchers among themselves use different and vague terms.

Eighth, merely to inform teachers about what constitutes "good" teacher behavior does not necessarily mean that teachers will change or even accept this as desirable. Many contend that teachers are conservative and are likely to resist changing their own behavior. Similarly, teachers lack the time to concentrate on modifying their teacher behavior.

Ninth, even when teachers seek feedback from researchers, especially from doctoral students who are conducting a study to complete their thesis, often, they are not provided with the results of such research, even though the findings can be duplicated and mailed to the principal or individual teachers. Many teachers lack understanding of research techniques, and they are unable to interpret findings, even if the researcher provides them with data. Other teachers seem to dismiss or resist research with "that is all good in theory, but it doesn't work in practice."

Tenth, many teachers seem no longer willing to cooperate with investigators from the colleges and universities. The reason is in part, since Sputnick there has been an increasing, near-compulsive disposition to criticize teachers. Since the "War on Poverty," this criticism seems to be focused on teachers of the

disadvantaged. The criticism seems unfair, wholesale, and flagrant, often couched in the angry rhetoric of exaggeration, and generally a biased presentation of one ghetto school (sometimes a few) from which the uncritical reader often proceeds to make generalizations about all ghetto schools and ghetto teachers. Many of these critics are popular writers who are unrelated to the colleges; nevertheless, in response to their widespread criticism, school teachers and administrators tend to generalize, too, that most outsiders are potential critics and a possible threat to the school or school system. Indeed, the lines of communications between school personnel and researchers seem strained, and this has serious implications for doctoral students and other researchers, who are often dependent on the teachers' and school officials' good will.

Conclusion

Lameke's (1955) comment of more than a decade ago bears repeating:

> If the research during the last three years were to be wiped out in the field of medicine, agriculture, physics, or chemistry, our lives would be materially changed. If research in the area of teacher personnel during the last three years would vanish, education and educators would continue as usual [p. 192].

It is sad but true that the most serious researchers of teacher behavior, a field closely related to teacher personnel, would probably not only make the same statement, but might extend it back to the turn of the century.

Ebel's (1967) criticism of research in general seems relevant to research on teacher behavior:

> Even today, when the prestige of science is at its height, most ... of the knowledge with which we guide our lives and solve our problems has come, not from controlled experiments, but from practical experience [p. 83].

> Most of the problems we face in the world today, including our educational problems, involve questions of purposes and values ... decisions that science could not possibly make for us [p. 83].

Most of the knowledge and acts which guide the teachers' behavior in the classroom are based, not on research, but personality, "common sense," and experience. Teaching involves an on-going interaction between teacher and students, problems arise that must be dealt with on the spot, as they occur - research does little good at that moment; since every situation is somewhat different, feelings, insights, hunches, etc. seem more important.

Concerning the limitations of theory in practical realms, Eisner (1963) cites Aristotle:

> ... it is the mark of an educated man to look for precision in each class of things just so far as the nature of the subject admits ... [p. 305]

If this viewpoint is translated into research on teacher behavior, it may suggest why the field is beset with so many problems and why so little worthwhile data have been forthcoming. The practical aspect of teacher behavior and teaching may simply be undefinable by research, or not capable of being subsumed by a research principle.

Finally, many researchers may find themselves agreeing with many of the above limitations and recommendations below of research on teacher behavior; they should note, however, the content of the paper is subjective and "armchair" in nature. No critical evaluation of the many studies and references was attempted even in the original chapter due to space limitations, and the interested reader - one who wishes to make his own approximate resolutions - will be obligated to read the materials on his own.

Recommendation for Future Research on Teacher Behavior

1. Granted, the status of research on teacher behavior is flaccid and imperfect; however, it can be improved, although it is not clear to what degree.

2. Rather than being trapped in analyzing teacher behavior, the investigators should first spend time in understanding its depth and complexity.

3. Investigators need to agree on (a) operational terms, (b) content of inventories, and (c) measurements of teacher behavior.

4a. Teacher-behavior terms should be formally defined, as are words in a dictionary.

4b. Teacher-behavior items should be neutral, that is, should take the same form regardless of subject, grade level, etc.

5. There is need to refine and formulate agreed-upon teacher behavior inventories. Terms and categories should be established in view of validity - content, concurrent, and predictive.

6. There should be agreement on measurement instruments, and on which instruments have equality of weights and units, beginning at the same point and preferably at zero.

7. Assumptions for using parametric tests, which are most popular among the researchers, should be indicated or at least made clearer. In some cases, where nonparametric tests should have been employed they were not; therefore, the findings are distorted.

8a. More attention is needed to understand the nature of teaching and the classroom process. Researchers tend to interpret data in terms of the behavorial sciences. There is need to translate findings into terms that are both (a) applicable to teaching and the classroom process, and (b) comprehensible to teachers.

8b. Under the guise of good scholarship, professors and researchers, especially in the field of education, tend to write for the benefit of their colleagues or other people in the field. The average reader is expected to know about what is being promulgated, but he often fails to comprehend it. This may be scholarly, but it does not help or affect most of the people who could benefit from "new" knowledge, i.e., teachers, themselves.

8c. In the meantime, teachers should try to understand research and apply findings to their own classroom situations.

9. Teaching involves a teacher-student interaction. Teacher-behavior research should be formulated in relationship to both teacher and student behavior. One without the other tends to be misleading and useless.

10. There is a need to learn to what extent teacher behavior is a function of personality.

11. In analyzing teacher behavior, it is important to consider the context in which it occurs. Much of the research, now, tends to treat teacher behavior as an isolated entity.

12. There is a need to control variables, at least the major ones, and to try to determine what kinds of teacher behaviors are desirable under what conditions and to what extent.

13a. There is a need to bring together and synthesize the numerous criteria, as well as the concepts and methods of conducting teacher-behavior research, into a framework which consists of a critical examination and comparisons, so that additional data can be hypothesized, developed, and analyzed in terms of previous data.

13b. Preference should be given to teacher-behavior criteria that have been commonly studied to maximize chances of validity and subsequent comparisons of data.

13c. Variables that cannot be agreed upon or organized should be tentatively discarded, for it would be easier, than, to validate teacher-behavior criteria, compare findings, and formulate hypotheses and theories.

14. Teacher behavior and teaching are often described as abstract, nebulous processes. There is a need to break these processes down into smaller and concrete components, which are recognizable and agreed upon, for purposes of analysis.

15a. A micro-analytical approach to the study of teacher behavior may be helpful, with well-defined criteria and agreed upon definitions, so that data may be more controllable, objective, and useful.

15b. Several micro-analytical studies might lead to teacher-behavior theories.

16. To the author's knowledge, very little, if any, research on teacher behavior has been conducted when teacher and students were using technological hardware in the classroom or school. Educational technology is increasingly used by teachers. Teacher-behavior research should move in this direction, since this seems to be a future trend with teaching and learning.

17. Investigators should take greater advantage of educational technology (computers, videotapes, records, etc.,) for purposes of facilitating, improving, and analyzing their research.

18. Much teacher-behavior research is conducted by doctoral students for purposes of earning a higher degree. University regulations, coupled with the candidate's desire to complete the study within approximately one year and his committee's desire to see him complete it, tend to make the candidate "play it safe" with a trivial problem. The idea is to get one's degree and then contribute something worthwhile to the field. Similarly, doctoral candidates usually lack one or more of the following: sufficient time, financial aid, staff assistance

or manpower, expertise, equipment, or facilities. This often leads to a somewhat worthless or useless study. For this reason, doctoral candidates should no longer be encouraged to conduct research on teacher behavior, unless it is a part of a more comprehensive study with sufficient funds, directed by an authority in the field.

19. Investigators, especially doctoral candidates, seem content to conduct their studies wherever they can find subjects. Teachers and students should be chosen by systematic selection and sampling, assuming they will cooperate, in order to make comparative studies.

20. Longtitudinal teacher-behavior studies should be conducted with agreement on criteria and variables.

21. Periodically, a nationwide teacher-behavior study might be conducted, noting socio-racial-geographical differences along with other agreed-upon variables, as a means for comparing other studies and teacher training programs.

22. Research on teacher behavior should be conducted in relation to the effects of various teacher-training programs.

23. We do not know how to train "good" teachers; we rely on descriptions, recommendations, and success stories; we basically use the same methods we were using at the turn of the century. This inability to train teachers becomes evident when teachers are assigned to work with the disadvantaged. The limited success of teachers seems linked more with personality than with training. Research on teacher behavior should be conducted in order to get away from the "technique," "story," "hit-or-miss" approach.

24. There is a need to improve rapport and communication between professors and teachers; moreover, investigators should provide clear and comprehensible feedback of their findings to teachers and school officials.

25. Criticism directed against teacher behavior and teaching, especially teachers who work with the disadvantaged, should cease, since no particular pattern of behavior can be advocated.

26a. In view of the growing demand for student power, especially at the college level, the value (or lack of value) of student evaluation of teaching performance should be judiciously considered. Do students have the right to evaluate faculty members? Are faculty members reasonably receptive to student evaluation? Do students' evaluations have a positive and continuing effect upon improving teaching (Academe, 1970)? What use will be made of the evaluations? These are serious questions, but what is more germane to our discussion is the reliability and validity of such evaluations. Several variables need to be considered, too. Space limitations permit the investigator to mention only a few: (1) the type of course -- required or elective; (2) size of student enrollment - a ten student seminar or a one hundred student lecture; (3) professor's distribution of grades; (4) degree of student unrest or dissent; (5) differences in students' and professor's social and political philosophy; (6) distinction between teacher behavior and course objectives; (7) professor's notoriety (a "halo effect") might affect the students' rating; (8) differences in the time (9 AM or 7PM), days (Monday or Saturday), or semesters (Fall or Summer); (9) professor's teaching load; (10) professor's access to secretarial or student assistance (ability to provide an abundance of mimeographed materials and quickly grade and return tests).

Until rating scales of teacher behavior are considered more reliable and valid, they should not be used for any purpose except for personal feedback. A "good" teacher or professor has little to worry about but can learn from his students. A "poor" instructor needs to know what students feel.

26b. Another trend that seems to be effecting the teaching profession, and especially directed toward inner-city schools, is the question of teacher "accountability." The question of who has the right and expertise to rate teachers may soon challenge the teacher probity and perhaps pitch teachers into conflict with students, parents, and/or supervisors. Similar questions, variables, and conclusions, as previously mentioned with regard to professors, also emerge.

26c. However, worthwhile and useful teacher-behavior ratings should enhance (1) the feasibility of determining merit pay and (2) the status of the teaching profession.

27a. Whether teachers are anti-research may no longer be a problem; in response to widespread criticism, they seem to be increasingly anti-researchers. This problem seems to be compounded by growing tension within the inner-city schools, whereby schools may be confronted with too many student-teacher-parent-administrator problems to risk having an outsider or researcher conduct a study or spotlight the school's plight - by reporting it in the study or publishing a commercial success. Similarly, in the context of the black power movement, northern schools that are controlled, in part, by black communities, or which have a large, militant black staff, will probably cease welcoming white investigators.

27b. In theory, almost any study conducted by a white investigator about the black community, school, or child can be construed as a potential Moynihan Report or Jensen exploration. Racial minorities and/or the educational establishment need not accept such findings, but they should acknowledge them. Hypersensitive egalitarianism may be the worst opponent of frank discussion and may impede racial equality. On the other hand, in view of the black-white

conflict, it is problematical whether educators should or could pursue or accept research findings that negatively depicts any racial or ethnic minority group. Until more social scientists are drawn from minority groups and are available to take part in research projects, the research establishment (including doctoral students who wish to study the disadvantaged) will probably be inhibited in exploring such areas.

REFERENCES

Berelson, B., & Steiner, G.A. Human behavior: An inventory of scientific findings. New York: Harcourt, 1964.

Bloom, B.S. (ed.), Englehart, M.D., Furst, E.J., Hill, W.H., & Krathwohl, D.R. Taxonomy of educational objectives, the classification of educational goals, handbook I: Cognitive domain. New York: McKay, 1956.

Broudy, H.W. Can we define good teaching. Record, 1969, 70, 583-592.

Dumas, W.W. Strengths and weaknesses of student teachers in English. Journal of Experimental Education, 1966, 35, 19-27.

Ebel, R. L. Some limitations of basic research in education. Phi Delta Kappan, 1967, 49, 81-84.

Eisner, E.W. Qualitative intelligence and the act of teaching. Elementary School Journal, 1963, 63, 299-307.

Flanders, N.A. Teacher influence, pupil attitudes, and achievement. Washington, D.C.: U.S. Department of Health, Education, and Welfare, 1965. Research Monograph No. 12.

Gallagher, J.J., & Aschner, M.J. A preliminary report on analysis of classroom interaction. Merrill-Palmer Quarterly, 1963, 9, 183-194.

Getzels, J.W., & Jackson, P.W. The teacher's personality and characteristics. In N.L. Gage (ed.), Handbook of research on teaching. Chicago: Rand McNally, 1963. Pp. 506-582

Goheen, R.F. The teacher in the university. American Scientist, 1966, 54, 221-225.

Guilford, J.P. Psychometric methods. (2nd ed.) New York: McGraw-Hill, 1954.

Guilford, J.P. Intelligence: 1965 model. American Psychologist, 1966, 21, 20-26.

Harvey, O.J., Prather, M., White, B.J., & Hoffmeister, J.K. Teachers' beliefs, classroom atmosphere and student behavior. American Educational Research Journal, 1968, 5, 151-166.

Hughes, M.M. What is teaching? One viewpoint. Educational Leadership, 1962, 19, 251-259.

Hughes, M.M., & Associates. Development of the means for the assessment of the quality of teaching in elementary schools. In M.D. Usdan & F. Bertolaet (eds.), Development of school-university programs for the pre-service education of teachers for the disadvantaged through teacher education centers. Washington, D.C.: U.S. Department of Health, Education, and Welfare, 1965, Pp. 159-167. Research Project No. F-068.

Jersild, A.T. Characteristics of teachers who are "liked best" and "disliked most." Journal of Experimental Education, 1940, 9, 139-151.

Lamke, T.A. Introduction. Review of Educational Research, 1955, 25, 192.

Masia, B.B. Evaluating educational outcomes by means of formal behavior science instruments. In M.D. Usdon & F. Bertolaet (eds.), Development of school-university programs for the pre-service education of teachers for the disadvantaged through teacher education centers. Washington, D.C.: U.S. Department of Health, Education, and Welfare, 1965, Pp. 214-251. Research Project No. F-068.

Medley, D.M. Teacher personality and teacher-pupil rapport. Journal of Teacher Education, 1961, 12, 152-156.

Meux, M., & Smith, B.O. Logical dimensions of teaching behavior. In B.J. Biddle & W. L. Ellena (eds.), Contemporary research on teacher effectiveness. New York: Holt, 1964. Pp. 127-164.

Ornstein, A.C. Selected teacher behavior attributes rated as desirable by ninth-grade disadvantaged students and ninth-grade teachers of the disadvantaged. Unpublished doctoral dissertation, New York University, 1970.

Perkins, H.V. A procedure for assessing the classroom behavior of students and teachers. American Educational Research Journal, 1964, 1, 249-260.

Remmers, H.H. Rating methods in research on teaching. In N.L. Gage (ed.), Handbook of research on teaching. Chicago: Rand McNally, 1963. Pp. 329-378.

Ryans, D.G. Characteristics of teachers. Washington, D.C.: American Council on Education, 1960.

Sontag, M. Attitudes toward education and perception of teacher behaviors. American Educational Research Journal, 1968, 5, 385-402.

Warren, P.B. A study of lower class and middle-upper class students' perceptions of the behavior traits of the effective teacher. Unpublished doctoral dissertation, New York University, 1968.

Zahorik, J.A. Classroom feedback of teachers. Journal of Educational Research, 1968, 62, 147-150.